

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2001-249856

(43)Date of publication of application : 14.09.2001

(51)Int.CI. G06F 13/00
G06F 15/177

(21)Application number : 2000-397033 (71)Applicant : INTERNATL BUSINESS MACH CORP <IBM>

(22)Date of filing : 27.12.2000 (72)Inventor : BARNETT BARRY STANLEY DOUGLAS CRAIG BOSSEN

(30)Priority

Priority number : 2000 478306 Priority date : 06.01.2000 Priority country : US

(54) METHOD FOR PROCESSING ERROR IN STORAGE AREA NETWORK(SAN) AND DATA PROCESSING SYSTEM

(57)Abstract:

PROBLEM TO BE SOLVED: To provide a method and a system for performing problem determination and fault isolation in a storage area network(SAN).

SOLUTION: The method to process an error in the storage area network(SAN) comprises a step to generate a SAN topology map, a step to generate a SAN problem determination information table(SPDIT) and a step to generate a SAN diagnostic table(SDT) by using the SAN topology map and the SPDIT as computer execution steps.

SAN のたのトボロジ・マップ		
CNE A	CNE B	CNE C
CAE A 監視 スイッチ→ポート CAE1	MONITORING MONITORING	MONITORING
CAE B 監視 スイッチ→ポート CAE2	MONITORING MONITORING	MONITORING
CAE C MONITORING	MONITORING MONITORING	MONITORING MONITORING

LEGAL STATUS

[Date of request for examination] 27.12.2000

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

ストレージ・エリア・ネットワーク (S A N) 内
でのエラー処理方法及びデータ処理システム 特開2001-249856

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2001-249856

(P2001-249856A)

(43) 公開日 平成13年9月14日 (2001.9.14)

(51) Int.Cl.⁷

G 0 6 F 13/00
15/177

識別記号

3 0 1
6 7 8

F I

G 0 6 F 13/00
15/177

テマコード(参考)

3 0 1 V
6 7 8 A

審査請求 有 請求項の数30 O L (全 20 頁)

(21) 出願番号 特願2000-397033(P2000-397033)
(22) 出願日 平成12年12月27日 (2000.12.27)
(31) 優先権主張番号 09/478306
(32) 優先日 平成12年1月6日 (2000.1.6)
(33) 優先権主張国 米国 (U S)

(71) 出願人 390009531
インターナショナル・ビジネス・マシーンズ・コーポレーション
INTERNATIONAL BUSINESS MACHINES CORPORATION
アメリカ合衆国10504、ニューヨーク州
アーモンク (番地なし)
(74) 代理人 100086243
弁理士 坂口 博 (外2名)

最終頁に続く

(54) 【発明の名称】 ストレージ・エリア・ネットワーク (S A N) 内でのエラー処理方法及びデータ処理システム

(57) 【要約】

【課題】 ストレージ・エリア・ネットワーク (S A N) で問題判別及び障害分離を行うための方法及びシステムを提供。

【解決手段】 ストレージ・エリア・ネットワーク (S A N) 内でのエラーを処理する方法は、コンピュータ実行ステップとしてS A Nトポロジ・マップを生成するステップと、S A N問題判別情報テーブル (S P D I T) を生成するステップと、S A Nトポロジ・マップ及びS P D I Tを用いてS A N診断テーブル (S D T) を生成するステップとを有する。

SANのためのトポロジ・マップ

	CAE A	CAE B	CAE C
CAE A	S P D I T/種類 情報 スイッチ・ポート CAE 1	(バスA→B)=(1→2) バス使用せず	(バスA→C)=(1→3)
CAE B	(バスB→A)=(2→1) バス使用せず	S P D I T/種類 情報 スイッチ・ポート CAE 2	(バスB→C)=(2→3)
CAE C	(バスC→A)=(3→1)	(バスC→B)=(3→2)	S P D I T/種類 情報 スイッチ・ポート CAE 3

ストレージ・エリア・ネットワーク（S A N）内 でのエラー処理方法及びデータ処理システム

特開2001-249856

【特許請求の範囲】

【請求項1】ストレージ・エリア・ネットワーク（S A N）内でのエラーを処理する方法であって、コンピュータ実行ステップとして、
S A Nトポロジ・マップを生成するステップと、
S A N問題判別情報テーブル（S P D I T）を生成するステップと、
前記S A Nトポロジ・マップ及びS P D I Tを用いてS A N診断テーブル（S D T）を生成するステップと、
を有することを特徴とするエラー処理方法。

【請求項2】前記S A Nトポロジ・マップはS A Nトポロジ・テーブルを有し、該S A Nトポロジ・テーブルの各々の行が通信アーキテクチャ要素（C A E）に対して一意的にマッピングされ、また前記S A Nトポロジ・テーブルの各々の列がC A Eに対して一意的にマッピングされており、C A Eはネットワーク・サービス・プロトコルを介して通信アーキテクチャ管理プログラム（C A M）によって首尾よく登録されたネットワーク接続装置であり、C A Mは前記S A Nのための問題判別（P D）機能を有し、またS A NのP D情報テーブル（S P D I T）を保持し、さらに前記通信アーキテクチャ（C A）は前記S P D I Tに格納された情報を伝達する能力を有するすべてのネットワーク接続要素を有することを特徴とする請求項1に記載のエラー処理方法。

【請求項3】前記S P D I Tは、前記C A上の各プロダクト又は要素に対応付けられた少なくとも一つのデータ・レコードを有することを特徴とする請求項2に記載のエラー処理方法。

【請求項4】前記C A上の各プロダクト又は要素に対応付けられた少なくとも一つのデータ・レコードは、プロダクト・ベンダ情報、プロダクト識別子情報、前記プロダクト又は要素によってサポートされた通信リンクの種類に関する情報、及び／又は前記プロダクト又は要素によって報告されるエラー情報の種類に関する情報からなる群から選択される1つ以上のデータ項目をさらに有することを特徴とする請求項3に記載のエラー処理方法。

【請求項5】前記エラー情報の種類は、前記プロダクト又は要素が拡張リンク・サービス（E L S）登録リンク事象レコード（R L I R）をサポートするかどうかを示すことを特徴とする請求項4に記載のエラー処理方法。

【請求項6】前記S D Tは、前記C A MによってC A Eから受け取ったエラーと前記S A Nトポロジ・マップからの情報とを格納することを特徴とする請求項4に記載のエラー処理方法。

【請求項7】ストレージ・エリア・ネットワーク（S A N）内のエラーを処理する方法であって、コンピュータ実行ステップとして、
通信アーキテクチャ管理プログラム（C A M）でのエラー・メッセージ受信するステップと、
リアルタイム診断アルゴリズム（R D A）を用いて前記

エラー・メッセージを処理するステップとを有し、
前記C A Mは前記S A Nのため問題判別（P D）機能を有し、該C A Mは前記S A NのP D情報テーブル（S P D I T）を保持し、さらに前記C A Mによって管理された通信アーキテクチャ（C A）は前記S P D I Tに格納された情報を伝達する能力を有するすべてのネットワーク接続要素を有することを特徴とするエラー処理方法。

【請求項8】前記C Aをサポートするネットワークは、
バンド内ファイバ・チャンネル通信リンクとバンド外通信リンクとを有することを特徴とする請求項7に記載のエラー処理方法。

【請求項9】前記S A Nは、前記ネットワークに接続した複数のストレージ装置と、
前記ネットワークに接続した複数のホスト・コンピュータとをさらに有し、
前記複数のホスト・コンピュータの少なくとも一つはC A Mを有することを特徴とする請求項7に記載のエラー処理方法。

【請求項10】S A Nトポロジ・マップを生成するステップと、
前記S A Nトポロジ・マップ及び前記S P D I Tを用いてS A N診断テーブル（S D T）を生成するステップとを、さらに有することを特徴とする請求項7に記載のエラー処理方法。

【請求項11】すでに前記C A Mによって受信され、かつ前記S D Tに格納された時間的に関連したエラー・メッセージを検索する一方で、時間的制約障害分離判別に対する時間的相関ウインドウ（T C W）の値を用いて前記受信エラー・メッセージを分析するステップと、

【請求項12】すでに前記C A Mによって受信され、かつ前記S D Tに格納された空間的に関連したエラー・メッセージを検索する一方で、空間的制約障害分離判別に対する空間的相関パス・データ構造（S C P）を用いて前記受信エラー・メッセージを分析するステップと、

【請求項13】をさらに有することを特徴とする請求項10に記載のエラー処理方法。

【請求項14】前記受信エラー・メッセージによって示されたエラーの種類に応じた重大度重みを用いて前記受信エラー・メッセージを分析するステップを、さらに有することを特徴とする請求項11に記載のエラー方法。

【請求項15】ストレージ・エリア・ネットワーク（S A N）においてエラー情報を伝達するためのデータ処理システムであって、
バンド内ファイバ通信リンクとバンド外通信リンクとを有し、かつ通信アーキテクチャ（C A）をサポートするネットワークと、

前記ネットワークに接続した複数のストレージ装置と、
前記ネットワークに接続した複数のホスト・コンピュータとを有し、さらに、
【請求項16】前記複数のホスト・コンピュータの少なくとも一つは、

問題判別（P D）機能を持つ通信アーキテクチャ管理プログラム（C A M）を融資、C A MはS A NのP D情報テーブル（S P D I T）を保持し、また前記C Aは前記S P D I Tに格納された情報を伝達する能力を有するすべてのネットワーク接続要素を有することを特徴とするデータ処理システム。

【請求項14】複数のC A Mをさらに有し、前記C AはプライマリC A Mと1つ以上のセカンダリC A Mとを有し、またセカンダリC A MはにプライマリC A Mに対して重複的に動作することを特徴とする請求項13に記載のデータ処理システム。

【請求項15】前記C Aは、一つ以上のC A要素（C A E）と一つ以上のC Aノン・パーティシパント（C A N）とをさらに有し、C A Eはネットワーク・サービス・プロトコルを介してC A Mによって首尾よく登録されたネットワーク接続装置であり、C A NはS A Nトポロジ発見プロセスを介して存在することがまだ知られていないC A Mによって登録されていないネットワーク接続装置であることを特徴とする請求項13に記載のデータ処理システム。

【請求項16】前記バンド内ファイバ・チャンネル通信リンク及びバンド外通信リンクは、単一の物理的通信リンクによって提供されることを特徴とする請求項13に記載のデータ処理システム。

【請求項17】ストレージ・エリア・ネットワーク（S A N）内のエラーを処理するためのデータ処理システムであって、

S A Nトポジ・マップを生成するための第1の生成手段と、

S A N問題判別情報テーブル（S P D I T）を生成するための第2の生成手段と、

前記S A Nトポジ・マップ及び前記S P D I Tを用い手S A N診断テーブル（S D T）を生成するための第3の生成手段とを、

有することを特徴とするデータ処理システム。

【請求項18】前記S A Nトポジ・マップはS A Nトポジ・テーブルを有し、該S A Nトポジ・テーブルの各々の行が通信アーキテクチャ要素（C A E）に対して一意的にマッピングされ、また前記S A Nトポジ・テーブルの各々の列がC A Eに対して一意的にマッピングされており、C A Eはネットワーク・サービス・プロトコルを介して通信アーキテクチャ管理プログラム（C A M）によって首尾よく登録されたネットワーク接続装置であり、C A Mは前記S A Nのための問題判別（P D）機能を有し、またS A NのP D情報テーブル（S P D I T）を保持し、さらに前記通信アーキテクチャ（C A）は前記S P D I Tに格納された情報を伝達する能力を有するすべてのネットワーク接続要素を有することを特徴とする請求項17に記載のデータ処理システム。

【請求項19】前記S P D I Tは、前記C A上の各プロ

ダクト又は要素に対応付けられた少なくとも一つのデータ・レコードを有することを特徴とする請求項18に記載のデータ処理システム。

【請求項20】前記C A上の各プロダクト又は要素に対応付けられた少なくとも一つのデータ・レコードは、プロダクト・ベンダ情報、プロダクト識別子情報、前記プロダクト又は要素によってサポートされた通信リンクの種類に関する情報、及び／又は前記プロダクト又は要素によって報告されるエラー情報の種類に関する情報からなる群から選択される1つ以上のデータ項目をさらに有することを特徴とする請求項19に記載のデータ処理システム。

【請求項21】前記エラー情報の種類は、前記プロダクト又は要素が拡張リンク・サービス（E L S）登録リンク事象レコード（R L I R）をサポートするかどうかを示すことを特徴とする請求項20に記載のデータ処理システム。

【請求項22】前記S D Tは、前記C A MによってC A Eから受け取ったエラーと前記S A Nトポロジ・マップからの情報を格納することを特徴とする請求項20に記載のデータ処理システム。

【請求項23】ストレージ・エリア・ネットワーク（S A N）内のエラーを処理するデータ処理システムであって、コンピュータ実行ステップとして、

25 通信アーキテクチャ管理プログラム（C A M）でのエラー・メッセージ受信する受信手段と、
リアルタイム診断アルゴリズム（R D A）を用いて前記エラー・メッセージを処理する処理手段とを有し、
前記C A Mは前記S A Nのため問題判別（P D）機能を有し、該C A Mは前記S A NのP D情報テーブル（S P D I T）を保持し、さらに前記C A Mによって管理された通信アーキテクチャ（C A）は前記S P D I Tに格納された情報を伝達する能力を有するすべてのネットワーク接続要素を有することを特徴とするデータ処理システム。

35 【請求項24】前記C Aをサポートするネットワークは、バンド内ファイバ・チャンネル通信リンクとバンド外通信リンクとを有することを特徴とする請求項23に記載のデータ処理システム。

40 【請求項25】前記S A Nは、前記ネットワークに接続した複数のストレージ装置と、
前記ネットワークに接続した複数のホスト・コンピュータとをさらに有し、
前記複数のホスト・コンピュータの少なくとも一つはC A Mを有することを特徴とする請求項23に記載のデータ処理システム。

45 【請求項26】S A Nトポジ・マップを生成する第1の生成手段と、
前記S A Nトポジ・マップ及び前記S P D I Tを用いてS A N診断テーブル（S D T）を生成する第2の生成

手段とを、さらに有することを特徴とする請求項23に記載のデータ処理システム。

【請求項27】すでに前記C A Mによって受信され、かつ前記S D Tに格納された時間的に関連したエラー・メッセージを検索する一方で、時間的制約障害分離判別に対する時間的相関ウィンドウ（T C W）の値を用いて前記受信エラー・メッセージを分析する第1の分析手段と、

すでに前記C A Mによって受信され、かつ前記S D Tに格納された空間的に関連したエラー・メッセージを検索する一方で、空間的制約障害分離判別に対する空間的相関パス・データ構造（S C P）を用いて前記受信エラー・メッセージを分析する第2の分析手段と、

をさらに有することを特徴とする請求項26に記載のデータ処理システム。

【請求項28】前記受信エラー・メッセージによって示されたエラーの種類に応じた重大度重みを用いて前記受信エラー・メッセージを分析する第3の分析手段を、さらに有することを特徴とする請求項27に記載のデータ処理システム。

【請求項29】ストレージ・エリア・ネットワーク（S A N）においてエラー情報を伝達するためのデータ処理システムで使用されるコンピュータ読み取り可能媒体のコンピュータ・プログラム・プロダクトであって、S A Nトポロジ・マップを生成する第1の命令と、S A N問題判別情報テーブル（S P D I T）を生成する第2の命令と、

前記S A Nトポロジ・マップ及びS P D I Tを用いてS A N診断テーブル（S D T）を生成する第3の命令と、を有することを特徴とするコンピュータ・プログラム・プロダクト。

【請求項30】ストレージ・エリア・ネットワーク（S A N）においてエラー情報を伝達するためのデータ処理システムで使用されるコンピュータ読み取り可能媒体のコンピュータ・プログラム・プロダクトであって、通信アーキテクチャ管理プログラム（C A M）でのエラー・メッセージ受信する第1の命令と、

リアルタイム診断アルゴリズム（R D A）を用いて前記エラー・メッセージを処理する第2の命令とを有し、前記C A Mは前記S A Nのため問題判別（P D）機能を有し、該C A Mは前記S A NのP D情報テーブル（S P D I T）を保持し、さらに前記C A Mによって管理された通信アーキテクチャ（C A）は前記S P D I Tに格納された情報を伝達する能力を有するすべてのネットワーク接続要素を有することを特徴とするコンピュータ・プログラム・プロダクト。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、改善されたデータ処理システム、特にコンピュータ・ネットワーク管理の

ための方法及び装置に関する。

【0002】

【関連する技術】ストレージ・エリア・ネットワーク（S A N）は、複数のホスト・コンピュータが複数のストレージ周辺装置を共用すること、特にファイバ・チャネル（F C）ネットワーク・スイッチを介してストレージ周辺装置を共用することを可能とする「オープン・システム」ストレージ・アーキテクチャである。F Cスイッチ、ホスト・システム、及びストレージ周辺装置は、異なるベンダによって製造されたものであってもよく、また異なるオペレーティング環境を有するものであってもよい。

【0003】現在、F C S A Nの端末間問題判別機能又は仕様が求められている。マルチ・ベンダ・システム、ネットワーク・スイッチ、及び周辺装置からなる複雑な構成によって、既存の二点間ストレージ構成よりもS A N環境下で問題判別を実行することが困難である。そのため、S A N環境で障害が生ずることで、システムの動作不能時間が増加するとともにシステムをメンテナンスするための費用も増加するであろう。

【0004】

【発明が解決しようとする課題】S A Nに接続し、かつ障害が生じているシステム及び／又は構成要素を識別する障害分離アルゴリズムを取り込む「オープン・システム」、リアルタイム、端末間、エラー検出アーキテクチャを定義する方法及び装置を持つことは有利であろう。

【0005】

【課題を解決するための手段】ストレージ・エリア・ネットワーク（S A N）で問題判別及び障害分離を行うための方法及びシステムを提供する。マルチ・ベンダ・ホスト・システム、F Cスイッチ、及びストレージ周辺装置からなる複雑な構成は、通信アーキテクチャ（C A）を介してS A Nに接続される。通信アーキテクチャ要素（C A E）は、ネットワーク・サービス・プロトコルを介してホスト・コンピュータ上の通信アーキテクチャ管理プログラム（C A M）によって首尾よく記録されたネットワーク接続装置である。C A Mは、S A N用の問題判別（P D）機能を有し、S A NのP D F情報テーブル（S P D I T）を保守する。C A Mは、S P D I Tに格納された情報の伝達を行うことが可能なすべてのネットワーク接続要素を有する。C A Mは、S A Nトポロジ・マップを利用し、またS P D I TはS A N診断テーブル（S D T）の生成に使用される。特定の装置において障害を持つ構成要素がエラーを生じ、該エラーによって同一ネットワーク接続パス上にある装置にもエラーが生ずる可能性がある。エラー・パケット又はエラー・メッセージをC A Mが受信するので、エラーはS D Tに格納され、各エラーをS D T内の他のエラーと時間的かつ空間的に比較することによって各エラーの分析が行われる。

もしC A Mがエラーを生成する候補であると判断され

ストレージ・エリア・ネットワーク (S A N) 内 でのエラー処理方法及びデータ処理システム

特開2001-249856

るならば、もし可能ならばCAFを交換するように報告される。

【0006】

【発明の実施の形態】図1は、本発明が適用されるデータ処理システムの構成を示す模式図である。コンピュータ100は、システム・ユニット110と、ビデオテープ示端末102と、キーボード104と、フロッピー・ドライブ及び他の種類の固定記憶媒体及び取り外し可能な記憶媒体を有するものであってもよいストレージ装置109と、マウス106とを備える。また、コンピュータ100は、さらに別の入力装置を備えることも可能である。コンピュータ100は、任意の適当なコンピュータ、例えばIBM (International Business Machines Corporation (Armonk, New York所在)) の製品であるAdvanced Interactive Executive (AIX)・オペレーティング・システムで稼働する同社のRISC/6000システムを用いて実現することができる。また、図示したコンピュータ100はサーバ型のコンピュータであり、本発明の他の実施形態を他のデータ処理システム、例えばワークステーション、ネットワーク・コンピュータ、ウェブ・ベース・テレビジョン・セットトップ・ボックス、インターネット機器等で実現してもよい。コンピュータ100は、該コンピュータ100で動作するコンピュータ読み取り可能媒体上にあるシステム・ソフトウェアによって実行可能なグラフィカル・ユーザ・インターフェースも含むものであってもよい。

【0007】なお、図1はあくまでも本発明の一例を示すものであって、本発明のアーキテクチャ的限界を示すものではない。

【0008】図2は、データ処理システムの内部構成要素の典型的な構成を説明するためのブロック図である。データ処理システム200は、多様なバス構造及びプロトコルを使用する。図に示した例ではPCIバス、ISAバス、及び6XXバスが用いられているが、他のバス・アーキテクチャ及びプロトコルを使用してもよい。

【0009】プロセッサ・カード201は、プロセッサ202及びL2キャッシュ203を有し、これらは6XXバス205に接続されている。システム200は、複数のプロセッサ・カードを有するものであってもよい。プロセッサ・カード206は、プロセッサ207及びL2キャッシュ208を有する。

【0010】6XXバス205はシステム・プレーナ210をサポートするもので、該システム・プレーナ210は、6XXプリッジ211と、メモリ・カード213をサポートするメモリ・コントローラ212とを有する。メモリ・カード213は、複数のデュアル・インライン・メモリ・モジュール (DIMM) 215及び216から構成されるローカル・メモリ214を有する。

【0011】6XXプリッジ211は、システム・バス222を介してPCIプリッジ220及び221に接続

する。PCIプリッジ220及び221は、種々のI/O構成要素及びインターフェースをサポートする固有のI/O (NIO) プレーナ223上に包含される。PCIプリッジ221は、ネットワーク・アダプタ224及びいくつかのカード・スロット225を通る外部データ・ストリームのための接続がPCIバス227を介して提供される。また、PCIプリッジ220は、PCIバス228を介して多様なI/O装置と接続する。ハード・ディスク229は、PCIバス218と接続するSCSIホスト・アダプタ230と接続してもよい。グラフィック・アダプタ231もまた図示したように直接又は間接的にPCIバス228に接続してもよい。

【0012】ISAプリッジ232は、PCIバス228を介してPCIプリッジ220と接続する。ISAプリッジ232はISAバス234を介したNIOコントローラ233との相互接続機能、例えばシリアル接続235及び236を提供する。フロッピー (登録商標)・ドライブ接続237は取り外し可能なストレージを提供する。キーボード接続238及びマウス接続239によってデータ処理システム200がユーザから入力データを受け取ることが可能となる。不揮発性RAM (NVRAM) 240は、システム破壊又はシステムエラー、例えば電力供給上の問題に対してある種のデータが消えないように保存しておく不揮発性メモリを提供する。システム・ファームウェア241もまたISAバス234と接続されており、初期BIOSを制御する。サービス・プロセッサ244は、ISAバス234と接続されており、システム診断又はシステム・サービスを行うための機能を提供する。

【0013】サービス・プロセッサ244はエラーを検出し、情報をオペレーティング・システムに渡す。エラーのものがエラーを検出した時点で合理的確実性で知られていても、あるいは知られていないてもよい。オペレーティング・システムは、ただエラーをログファイルに書いてもよく、さもなければ報告されたエラーを処理してもよい。

【0014】当業者は、図2に示すハードウェアがシステム・インプリメンテーションに応じて変わることを理解するであろう。例えば、システムはより多くのプロセッサを有するものであってもよく、また他の周辺機器を図2に示したハードウェアに加えて、あるいは交換して使用してもよい。図示した例は本発明に関してアーキテクチャの限界を意味するものではない。

【0015】ここで図3を参照する。この図は、本発明の好ましい実施の形態にもとづいて実現されたSAN問題判別方法論に参加するデータ処理システムのための通信アーキテクチャを示す。ネットワーク300は、一組のコンピュータ、スイッチ、及びストレージ装置を有するもので、ストレージ装置は通信アーキテクチャに参加するか、又は参加しなくてもよい。

ストレージ・エリア・ネットワーク（S A N）内 でのエラー処理方法及びデータ処理システム

特開2001-249856

【0016】通信アーキテクチャ（CA）は、以下に詳細に説明されるS A N問題判別情報テーブル（S P D I T）に定義された情報のいずれか、又はすべてを伝達することが可能なS A N接続要素のすべてを含む。

【0017】CAに参加している各S A N接続要素は、CA要素（C A E）と呼ばれる。CAに参加していない要素はいずれもCA不参加者（C A N）と呼ばれる。これらの要素は識別される。なぜなら、これらの要素はS A Nトポロジに参加しており、それによってシステムの問題判別（P D）機能にも参加している。WindowsNT（商標）サーバ302、メインフレーム・コンピュータ304、Unix（商標）サーバ306、Linux（商標）サーバ308は、CAに参加するコンピュータであり、したがってこれらはC A Eである。WindowsNT（商標）サーバ302、メインフレーム・コンピュータ304、Unix（商標）サーバ306、Linux（商標）サーバ308は、様々なクライアントをサポートするホスト・コンピュータでもあり、ストレージ装置へのアクセスを要求してもよい。コンピュータ302～306の各々は、ホスト・バス・アタッチ（H B A）を有するもので、該H B AはF Cホスト用のネットワーク・アダプタの一種である。F Cスイッチ311～313はC A Eであり、ストレージ装置のいくつかもC A Eである。この例では、共有R A I D（独立したディスクの重複アレイ）321～323及び共有テープ324がC A Eであり、一方共有テープ325はC A Nである。

【0018】CAは、T C P / I Pプロトコルを用いるバンド内通信リンク341～352及び／又はすべてがS A N要素共有である通信リンク331～334上のバンド外T C P / I P通信ネットワークを介するF Cスイッチング・ファブリックを介して通信することができる。ここで指摘しておくべきことは、図3に示した通信リンクは単一の物理的接続を共有する論理接続であってもよいことである。あるいは、複数の物理的通信リンクによって装置が接続されてもよい。

【0019】情報を発行及び／又は収集するためにCAによって使用されたプロトコルは、S N M P / M I B（シンプル・ネットワーク・マネジメント・プロトコル／マネージメント・インフォメーション・ベース、モニタされている特定の装置を記述するS N M P構造）及びもととなった固有F Cの両方となるように定義される。これら2つのプロトコルを使用することで、装置／ホスト特異的及びS A N特異的情報の両方を収集する事が可能となり、次いで端末間問題判別に使用される。

【0020】CA管理プログラム（C A M）は、システムの端末間P D機能が存在する特定のC A Eである。S P D I TはC A Mに存在し、C A EごとにC A Mによって自動的に登録される（固有F C及び／又はS N M Pサービスを介して）。C A Eは、首尾よく登録する要素であり、またC A NはC A Mによって登録できない要素で

はあるがS A Nトポロジ発見プロセスを介して存在することが知られている。このことについては後でより詳細に説明する。C A Mは、端末間問題判別に関する任意のF C拡張リンク・サービス（E L S）をサポートする。

【0021】C A Mは、プライマリ又はアクティブC A M、及びセカンダリ又はイナクティブC A Mとして分類することができる。C A Mは、S P D I T及び登録情報を複製する大いに利用可能な要素である。例えば、セカンドリC A M及びプライマリC A Mは、もしプライマリC A Mがハートビート信号に応答しないことによって失敗したように思われるならば、重複する方法で実行されているセカンダリC A MがプライマリC A Mのデュアルを想定するように、ハートビート信号を共有してもよい。C A Mに対する問題判別インターフェースは、S A N P Dアプリケーション・プログラミング・インターフェース（S A N P D A P I）から構成される。S A N P D A P Iは、C A Mと、C A M情報又はステータスを読むことが可能な任意の他のオペレーティング環境との間の通信インターフェースを定義する。

【0022】ここで図4を参照する。本発明の好ましい実施の形態にもとづくS A N問題判別情報テーブル（S P D I T）を示すテーブルである。S P D I Tは、すべての周知のプロダクト／要素とC A 上で伝達可能な情報種類とで構成されている。S P D I Tのフォーマットは、C A に含まれる装置の数、サポートされるプロダクトの種類、装置に関連した情報、その他にかなり依存していると思われる。例えば、S P D I Tは、図3に示す各々の装置に関する情報を含むであろう。

【0023】S P D I T 4 0 0は、以下のレコード項目を含むものであってもよい。すなわち、ベンダ属性401、プロダクト識別子402、情報の種類403、及び記述属性404である。S P D I T 4 0 0の各レコードは、それらのレコード項目に関するデータを含む。ベンダ属性401は、C A 上の特定装置製造者を含む。プロダクト識別子402は、特定の装置を識別するためのベンダによって割り当てられた情報、例えば型の種類、型番号、プロダクト通し番号等である。

【0024】情報の種類403は、装置によってサポートされた通信リンクの種類、装置によってサポートされたエラー条件又はエラー定義のフォーマット等に関連したデータを含む。記述属性404は、プロダクトによって予想されるべきエラー情報の種類に関する情報を提供する。例えば、もし記述属性レコードが、プロダクトはE L S登録リンク発生事象レコード（R L I R）互換性であることのテーブル示のみを含むならば、C A M関連プロセスはプロダクトのバンド外M I Bを受け取ることは期待できないであろう。

【0025】S P D I Tは、一般にS A N対応周辺機器、ホスト、及びスイッチによって状態／エラー条件を

示すことに使われる情報のすべてを含むであろう。これは、固有FCリンク及び拡張リンクエラー定義、さらにMIB定義を含むと思われる。これらの定義は、フィールド置換可能ユニット(FRU)構成要素情報を含むことができ、MIBに位置するか、又はエラー報告プロトコルに組み込まれることができ、また障害構成要素が分離される細分性の決定に使用可能である。

【0026】すでに指摘したように、CAMはシステムの端末機能が存在する特有のCAであり、SPDITが含まれる。CAM初期化プロセスは、バンド内SAN及びバンド外ネットワークの両方に接続したすべてのFCノードの発見及び登録を含む。CAM初期化プロセスは、FCバンド内及びCAバンド外(SNMPを経由)発見/登録プロセスを使用する。このプロセスは、SANのトポロジ・マップ(TM)を提供する。このトポロジ・マップ(TM)は、要素の種類(ホスト、周辺機器、スイッチ)、明示接続/バス、及びそれらの関連ベンダの知識、さらにSPDIT情報を伴って要素が接続した登録及び非登録SANのすべてを含む。

【0027】ここで図5を参照する。この図では、SANのための簡略化ネットワーク・トポロジが図示されている。FCスイッチ501は、該FCスイッチ501とCAE521~523の各々との接続点を与えるポート511~513を有するもので、該ポート511~513はCAE A、CAE B、及びCAE Cとラベルされている。CAの見地からすれば、FCスイッチポート511~513はCAEである。なぜなら、これらのポート511~513は障害を生じたり、又はエラーを生じたりすることが可能であり、エラーのもとと適当に診断された後は取り替えることができる。

【0028】図6は、図5に示すSANのための簡略化ネットワーク・トポロジのマップを提供するテーブルである。トポロジ・マップ(TM)は2次元のテーブルとしてテーブルされており、左側の欄と上側の欄との両方に、図5のFCスイッチ501のようなスイッチに接続されたSAN要素、CAE及びCAN装置の両方が含まれている。診断セルは、対応の要素及びそれが接続されたスイッチ・ポートに関するSPDIT/種類情報のすべてを含む。他のセルは要素間の方向バスが含まれる。例えば、テーブルはポート3とポート1との間のバスを用いてCAEAとCAE Cとの間の方向バスを示している。多重バスは可能である。トポロジ及び登録発見プロセスは周期的に繰り返されることでTMが正しいことを確実にする。CAMもまた、PFのために使用可能な拡張リンク・サービスを提供する任意のSAN要素によって登録する。

【0029】図7は、SAMのためのSAN診断テーブルを示すものである。SANのTMは、第1エラー・データ収集(FEDC)及びリアルタイム診断分析(RDA)のために使用されるSAN診断テーブル(SDT)

を生成するために使われる。図7に示されるSDTは、各スイッチ/ファブリック要素のための追加の行が含まれること以外は、図6に示すTMと同様である。

- 【0030】対角線上のSDTセルは、スイッチ・ポートが含まれるそれ自身の行/列に対応するCAEによって報告されたエラーを保持するために使われる。一つのバス、すなわちSDTセルの中の各点は、別のSAN接続要素をテーブルす。各セルは、特定のプロダクト動作が知られ、かつ適当な診断決定が下されるように、TMに収集された情報を含む。例えば、もしバンド外SCSI装置のエラーがバンド内ホスト・バス・アタッチ(HBA)FCリンクのエラーを伴うならば、ストレージ装置はHBAに対してリンクエラーを生じている可能性が高いかどうか等の診断問い合わせが存在してもよい。
- 【0031】図7に含まれる模範的なエラー情報は、SDTを用いるRDAのユーティリティを例証するものである。行1は、バンド内FCリンクがタイムアウトであることをCAE Aが報告したことを示している。行3は、CAE C上のバンド外ハードウェア・コントローラエラーを示す。これら2つのエラーは関連している。なぜなら、エラー情報に関連したタイムスタンプによって示されるように、これらは同一タイム・フレーム内で起こるためである。行5はバンド内FCリンクエラーが生じたことを示すが、記憶されたタイムスタンプという条件のもとで、行5は前の2つとは無関係である。したがって、テーブルは2つの異なる問題を示している。すなわち、第1の問題は、CAE Cにおける制御装置のハードウェア障害に関係しており、第2の問題はFCスイッチのCAE 2におけるFCリンク障害である。
- 【0032】図8は、種々のエラーに対するリアルタイム診断分析で使用される重みを図示している。RDAアルゴリズムは、FEDC事象が起こるたびに、SDTのエラー報告要素を走査して適当な応答を割り出す。RDAは、障害構成要素を分離するために重み判断分析を使用する。2つの広範囲のカテゴリーがH=最も高い重み、M=中間の重み、L=最も低い重みによって例証される。
- 【0033】SDT走査アルゴリズム及びエラー重みはダイナミックであり、SANトポロジの複雑性及びそれに接続した要素の性質に合わせて帰られるであろう。
- 【0034】図に示す重みテーブルは、典型的なSAN環境に適用される強弱間の重みスケールを簡単に例証するものである。もしSANがその対応したホスト及び周辺機器でほんのわずかな16ポート・スイッチまで成長するならば、単一のドライブエラー又はHBAタイムアウトエラーによるエラーを報告することができる可能なノード数を増やすことができる。グローバルな端末間RDA診断機能無しで、障害構成要素を分離するタスクは行き当たりばったりになる。マルチ・ベンダSANでは、多数の断続的な、回復可能な装置エラー(すなわち

ソフトのエラー) がホストによって気づかれなくなることは一般的である。結局、装置は回復不能エラー(すなわちハード・エラー)に遭遇する可能性があり、それはシステムクラッシュをもたらす。本発明によって提供されるバンド内及びバンド外機構は、回復可能なエラーが生ずるとただちにそれを検出して報告するであろう。

【0035】ここで、図8ないし図12を参照しながら説明する。これらの図は、本発明の好ましい実施の形態にもとづく単一障害S A N要素のS A N端末間障害分離を行うリアルタイム診断アルゴリズム(R D A)のプロセスを説明するためのフローチャートである。T D Aは、障害を分離するために2つのダイナミックな機構を使用する。

【0036】1. 時間相関関係ウィンドウ(T C W)
T C Wはスカラーの値であり、すなわち時間の範囲、誤診の可能性が時間次元で最小となるように該時間次元においてS D Tの障害分離検索を制限するために使用される。

【0037】2. 空間相関関係パス(S C P)
S C Pはデータ構造であって、誤診の可能性が空間次元で最小となるように、また周知のシステムとサブシステムとの関連が綿密に調べられるように、該データ構造はS D Tの空間ドメインにおける障害分離検索を制限するために使用される。S C PはR D Aの間、S D Tから要素をコピーする。

【0038】R D Aのゴールはかなりの確実性で障害のもとが分離されるまで、時間、場所、及び重大度について、受け取った障害情報のすべてを関連づけることである。このプロセスは、单一の報告された障害、又は一連の報告された障害の後に終了する。

【0039】単一障害S A N要素のS A N端末間障害分離用の一般的なR D Aを以下に説明する。C A Mが該C Aにあるすべての接続バスをS D Tに初期化する時に、プロセスが開始する(ステップ801)。S D Tは、すべての接続されたバス、すなわちバスA-->B、B-->C、その他によって初期化される。接続可能とすべきそれらのバスだけが入力される。これらのバスはトポロジ・マッピング、例えば図5及び図6に示したT Mと同様のT Mによって確立される。S A Nは、接続可能とすべきでないある特定の接続から外すために、完全には接続しない可能性がある。例えば、あるホストは特定のストレージ装置に対するデータの格納及び検索を制限されるかもしれない。システム管理者は、メインフレームがN Tデータを破損又は破壊する能力を持たないよう、N Tホストがデータを特定の装置に格納することを可能とすることができます。

【0040】プロセスは、S A NのためにT C W及びS C Pを初期化することで継続する(ステップ802)。T C Wは、時間ウィンドウであり、また時間の値、通常は秒ないし分のオーダーで要求する。S C Pは、S D T

から選択されたバスのすべてのセットを含んでいる。これらのバスは、周知のホストとストレージ装置との間、ホストとホストの間、ストレージ装置とストレージ装置との間の関係を反映するもので、該関係はトポロジ・マッピングによって確立される。再びここで指摘しておくべきことは、セカンダリC A MはプライマリC A Mに格納されているデータ構造及び値を保全することである。

【0041】続いて、C A Mは新たなエラーを受け取り(ステップ803)、R D Aを用いてエラーを処理する(ステップ804)。R D Aプロセスが終了したかどうかについての判断を行い(ステップ805)、もし終了していなければ、プロセスはステップ803に戻ってさらにエラーを受け取り、かつ処理する。もし終了していれば、S A N端末間障害分離のための初期化プロセスが完了する。

【0042】図10は、新しいエラーの処理ステップ、例えば図9のステップ804をさらに詳細に説明するフローチャートである。プロセスは、新しいエラーを受け取ることによって開始し(ステップ810)、エラー、エラーが生じた時間、及びエラーの重大度(高、中、低)を報告する構成要素を示すためにS D Tが更新される(ステップ811)。エラーが重大度の高いものであるかどうかについて判断を行う(ステップ812)。もし重大度が高ければ、このエラーはメンテナンスを必要とする障害であると直ちに報告される(ステップ813)。続いて、報告されたエラーが交換すべき特定の部品に関連するかどうかについて判断するために、S P D I Tが問い合わせされる(ステップ814)。もし関連していなければ、高重大度のエラーの処理が完了する。もし関連していれば、障害構成要素の交換が指示され(ステップ815)、高重大度のエラーの処理が完了する。

【0043】もしエラーが高重大度エラーでなければ、該エラーが中又は低度の重大度からなるエラーであるかどうかについて判断される(ステップ816)。もしそうであるならば、低／中重大度エラーが処理され(ステップ817)、さらに該エラー処理が完了する。

【0044】もし、エラーが高重大度エラーでも低／中重大度エラーでもなければ、エラーの重大度は誤りとして判断され、該エラーが無視される(ステップ818)。

【0045】ここで、図11を参照しながら説明する。この図では、図10のステップ817のように、新たな低／中重大度エラーの処理を行う方法をより詳細に説明するためのフローチャートが示されている。S C Pは報告されたエラーによって影響を受けるバスを判断するのに用いられる。これらのバスにおいて要素のS D Tセルの各々は、新しいエラーを報告している要素を含むもので、前のエラー発生に対して順番に問い合わせられる(ステップ820)。そして、前のエラーの発生が現在

のエラーと空間的に関係しているかどうかについて判断される（ステップ821）。そして、問い合わせは前のエラーが現在のエラーと空間と同様に時間的に関係しているかどうかについて判断するためにT C Wを使用する（ステップ822）。もし、前のエラーが時間的かつ空間的に関係しているならば、エラーはS C Pに格納される（ステップ823）。問い合わせが終了すると、S C Pは、時間的制約のなかで生じたS D Rの適当なバス上でのすべてのエラーのマッピングが含まれる。

【0046】S C Pのデータ構造がS C Pのためのデータ構造が組織化され、かつ使用される方法は、システム・インプリメンテーションに応じて変化するものであってもよい。例えば、S D Tからの要素は、S C Pにコピーされ、データは空間又は時間に関連していないと判断させるので、要素をS C Pからデリートすることができる。

【0047】アルゴリズムは、障害構成要素の位置を分離するためにエラー相関関係／重大度の評価をしなければならない。ここで図12を参照する。この図では、低／中重大度のエラーに対応づけられた障害構成要素についていくつかの可能なケースを説明するフローチャートが示されている。

【0048】プロセスは、新たに受信されたエラーを生成した現在の要素からすべてのエラーが生ずるかどうかについて判断することから開始する（ステップ830）。もしそうならば、2つ以上のエラーがS C Pに存在するかどうかについての判断が行われる（ステップ831）。もしうでなければ、現在のエラーの処理が完了する。もしうであるならば、現在の要素はメンテナンスを必要とすることが示される（ステップ832）。次にS P D I Tは、報告されたエラーが交換すべき特定の部品と対応づけられているかどうかについて判断するために問い合わせられる（ステップ833）。もしうであるならば、障害構成要素の交換が指示され（ステップ834）、新たに受信され、低／中重度大エラーが処理される。

【0049】もしすべてのエラーが現在の要素から生ずるならば、すべての（2つの以上）エラーが单一のバスに含まれるかどうかについて判断がなされる（ステップ835）。この場合、バス内の任意の要素が報告されたエラーの根本原因であるかもしれない、また装置ハードウェア関連のエラーはリンク又はタイムアウト関連のエラーに優先される。エラーに装置ハードウェアエラーが含まれるかどうかについて判断がなされる（ステップ836）。もし装置ハードウェアが見いだされるならば、ステップ832～834と同様にして、関連した要素がメンテナンスを必要とすることが示される。そして、S P D I Tは報告されたエラーが交換すべき特定の部品に関連づけられているかどうかについて判断するように問い合わせられる。もしうであるならば、障害構成要素の

交換が指示される。

【0050】もし信号バス上のエラーに装置ハードウェア・エラーが含まれていなければ、リンク又はタイムアウト・エラーのみが報告されている。この状況によつ

- 05 て、リンクの性能の劣化と最終的な障害が導かれる。このような場合、アルゴリズムは最初にエラーを報告している要素を探す（ステップ837）。すなわち、最初のエラーの優先度が高く、また他のものは最初の発生と関係があると考えられる。複数のエラーが連鎖的に生じて10 いる要素がひとたび見つかると、上記したステップ832～834と同様に、関連した要素がメンテナンスを必要とすることが示される。そして、S P D I Tは報告されたエラーが交換すべき特定の部品に関連づけられているかどうかについて判断するよう問い合わせられる。も15 しそうであるならば、障害構成要素の交換が指示される。

【0051】もし2つ以上のエラーが单一のバスに含まれていなければ、2つ以上のエラーが多数のバスに生じている。多数のエラーが生じている複数のバス上に任意の共通の要素が存在するかどうかを判断する（ステップ838）。もし存在するならば、この場合はそれらのバス上の共通要素分離と（ステップ839）、エラー相関／重大度の評価の実行とが必要となる。

- 【0052】共通要素は、S A N終端要素及び／又はS 25 ANファブリック要素のいずれか一方であることが可能である。S A N終端又はファブリック要素が唯一の共通要素であるかどうかの判断が行われる（ステップ840）。もしうであるならば、ステップ832～834と同様に、この共通要素の障害があるものとして示され30 れ、メンテナンスが必要とされる。そして、S P D I Tは報告されたエラーが交換すべき特定の部品に関連づけられているかどうかについて判断するよう問い合わせされる。もしうであるならば、障害構成要素の交換が指示される。

- 【0053】さもなければ、もしS A N終端又はファブリック要素が唯一の共通要素でなければ、S A N終端及びS A Nファブリック要素の両方が共通要素である。この状況は、今やステップ835での判断の結果に等しく、プロセスはさらに処理を行うためにステップ836 40 に分岐する。

- 【0054】もし単一バスに含まれていない2つ以上のエラーがあり、またこれらのエラーに共通の要素が存在しないならば、多数のエラーの各々に対して別々にリアルタイム診断アルゴリズム（R D A）が実行される（ステップ841）。このことは希ではなるが、実現可能なシリオがT C Wに一つ以上のエラーが受信され、該エラーが別々に障害構成要素が生じる場合に起こりうる。この点で、あたかも各エラーが新しく受信されたエラーであるかのように各エラーを処理するために、エラー・プロセスはステップ804に戻る。

ストレージ・エリア・ネットワーク (S A N) 内 でのエラー処理方法及びデータ処理システム

特開2001-249856

【0055】本発明の利点は、上述した本発明の詳細な説明に鑑みて明らかである。S A N診断テーブルはS A Nトポロジ、固有のファイバ・チャンネル・サービス、及びベンダ特定情報を用いて生成される。本発明は、S A N問題判別のためにF C固有バンド内及びホスト／装置特異的バンド外状況／エラー・データ収集の両方をサポートする。そして、リアルタイム診断アルゴリズムはS A N診断テーブルをトラバースして、障害S A N構成要素を分離する。この方法論は有利である。なぜなら、この方法論は管理端末に対する固有のアクセス又は装置診断が障害構成要素を分離する上で必要とされないようにホストのオペレーティング環境に実装することが可能である。さらに、この方法論は、プラットフォームに対して独立しており、またS A N問題判別のためにF C固有バンド内及びホスト／装置特異的バンド外状況／エラー・データ収集の両方をサポートする。

【0056】ここで指摘しておくべき重要なことは、完全に機能しているデータ処理システムというかたちで本発明を説明した一方で、当業者が本発明のプロセスがコンピュータが読み取り可能な媒体の形態及び種々の形態で分配可能であること、また本発明は分配を実行するために実際に使用される信号保持媒体の特定の種類に関わりなく等しく適用されることである。コンピュータが読み取り可能な媒体の例として、フロッピー・ディスク、ハード・ディスク・ドライブ、RAM、及びCD-R O M等の記録可能型媒体、デジタル及びアナログ通信リンク等の転送型媒体が挙げられる。

【0057】以上、本発明の詳細な説明を本発明の例証及び説明を目的として行ったが、記述された形態で本発明を網羅及び限定することを意図したものではない。当業者は本発明の多くの改良例及び変形例を容易に想到することができよう。また、本発明の原理、実質的な用途を最良のかたちで説明するために、また、発明の詳細な説明では、特定の使用形態に適するように様々な修飾が施された様々な実施の形態を当業者が理解可能となるように、実施の形態が選択されて記述されている。

【0058】まとめとして、本発明の構成に関して以下の事項を開示する。

(1) ストレージ・エリア・ネットワーク (S A N) 内でのエラーを処理する方法であって、コンピュータ実行ステップとして、S A Nトポロジ・マップを生成するステップと、S A N問題判別情報テーブル (S P D I T) を生成するステップと、前記S A Nトポロジ・マップ及びS P D I Tを用いてS A N診断テーブル (S D T) を生成するステップと、を有することを特徴とするエラー処理方法。

(2) 前記S A Nトポロジ・マップはS A Nトポロジ・テーブルを有し、該S A Nトポロジ・テーブルの各々の行為が通信アーキテクチャ要素 (C A E) に対して一意的にマッピングされ、また前記S A Nトポロジ・テーブル

の各々の列がC A Eに対して一意的にマッピングされており、C A Eはネットワーク・サービス・プロトコルを介して通信アーキテクチャ管理プログラム (C A M) によって首尾よく登録されたネットワーク接続装置であ

り、C A Mは前記S A Nのための問題判別 (P D) 機能を有し、またS A NのP D情報テーブル (S P D I T) を保持し、さらに前記通信アーキテクチャ (C A) は前記S P D I Tに格納された情報を伝達する能力を有するすべてのネットワーク接続要素を有することを特徴とする上記 (1) に記載のエラー処理方法。

(3) 前記S P D I Tは、前記C A上の各プロダクト又は要素に対応付けられた少なくとも一つのデータ・レコードを有することを特徴とする上記 (2) に記載のエラー処理方法。

(4) 前記C A上の各プロダクト又は要素に対応付けられた少なくとも一つのデータ・レコードは、プロダクト・ベンダ情報、プロダクト識別子情報、前記プロダクト又は要素によってサポートされた通信リンクの種類に関する情報、及び／又は前記プロダクト又は要素によって報告されるエラー情報の種類に関する情報からなる群から選択される1つ以上のデータ項目をさらに有することを特徴とする上記 (3) に記載のエラー処理方法。

(5) 前記エラー情報の種類は、前記プロダクト又は要素が拡張リンク・サービス (E L S) 登録リンク事象レコード (R L I R) をサポートするかどうかを示すことを特徴とする上記 (4) に記載のエラー処理方法。

(6) 前記S D Tは、前記C A MによってC A Eから受け取ったエラーと前記S A Nトポロジ・マップからの情報とを格納することを特徴とする上記 (4) に記載のエラー処理方法。

(7) ストレージ・エリア・ネットワーク (S A N) 内のエラーを処理する方法であって、コンピュータ実行ステップとして、通信アーキテクチャ管理プログラム (C A M) でのエラー・メッセージ受信するステップと、リアルタイム診断アルゴリズム (R D A) を用いて前記エラー・メッセージを処理するステップとを有し、前記C A Mは前記S A Nのため問題判別 (P D) 機能を有し、該C A Mは前記S A NのP D情報テーブル (S P D I T) を保持し、さらに前記C A Mによって管理された通信アーキテクチャ (C A) は前記S P D I Tに格納された情報を伝達する能力を有するすべてのネットワーク接続要素を有することを特徴とするエラー処理方法。

(8) 前記C Aをサポートするネットワークは、バンド内ファイバ・チャンネル通信リンクとバンド外通信リンクとを有することを特徴とする上記 (7) に記載のエラー処理方法。

(9) 前記S A Nは、前記ネットワークに接続した複数のストレージ装置と、前記ネットワークに接続した複数のホスト・コンピュータとをさらに有し、前記複数のホスト・コンピュータの少なくとも一つはC A Mを有する

ストレージ・エリア・ネットワーク (SAN) 内 でのエラー処理方法及びデータ処理システム

特開2001-249856

ことを特徴とする上記(7)に記載のエラー処理方法。

(10) SANトポロジ・マップを生成するステップと、前記SANトポロジ・マップ及び前記SPDITを用いてSAN診断テーブル(SDT)を生成するステップとを、さらに有することを特徴とする上記(7)に記載のエラー処理方法。

(11) すでに前記CAMによって受信され、かつ前記SDTに格納された時間的に関連したエラー・メッセージを検索する一方で、時間的制約障害分離判別に対する時間的相関ウィンドウ(TCW)の値を用いて前記受信エラー・メッセージを分析するステップと、すでに前記CAMによって受信され、かつ前記SDTに格納された空間的に関連したエラー・メッセージを検索する一方で、空間的制約障害分離判別に対する空間的相関パス・データ構造(SCP)を用いて前記受信エラー・メッセージを分析するステップと、をさらに有することを特徴とする上記(10)に記載のエラー処理方法。

(12) 前記受信エラー・メッセージによって示されたエラーの種類に応じた重大度重みを用いて前記受信エラー・メッセージを分析するステップを、さらに有することを特徴とする上記(11)に記載のエラー方法。

(13) ストレージ・エリア・ネットワーク(SAN)においてエラー情報を伝達するためのデータ処理システムであって、バンド内ファイバ通信リンクとバンド外通信リンクとを有し、かつ通信アーキテクチャ(CA)をサポートするネットワークと、前記ネットワークに接続した複数のストレージ装置と、前記ネットワークに接続した複数のホスト・コンピュータとを有し、さらに、前記複数のホスト・コンピュータの少なくとも一つは、問題判別(PD)機能を持つ通信アーキテクチャ管理プログラム(CAM)を融資、CAMはSANのPD情報テーブル(SPDIT)を保持し、また前記CAは前記SPDITに格納された情報を伝達する能力を有するすべてのネットワーク接続要素を有することを特徴とするデータ処理システム。

(14) 複数のCAMをさらに有し、前記CAはプライマリCAMと1つ以上のセカンダリCAMとを有し、またセカンダリCAMはにプライマリCAMに対して重複的に動作することを特徴とする上記(13)に記載のデータ処理システム。

(15) 前記CAは、一つ以上のCA要素(CAE)と一つ以上のCANON・パーティシパント(CAN)とをさらに有し、CAEはネットワーク・サービス・プロトコルを介してCAMによって首尾よく登録されたネットワーク接続装置であり、CANはSANトポロジ発見プロセスを介して存在することがまだ知られていないCAMによって登録されていないネットワーク接続装置であることを特徴とする上記(13)に記載のデータ処理システム。

(16) 前記バンド内ファイバ・チャンネル通信リンク

及びバンド外通信リンクは、単一の物理的通信リンクによって提供されることを特徴とする上記(13)に記載のデータ処理システム。

(17) ストレージ・エリア・ネットワーク(SAN)
05 内のエラーを処理するためのデータ処理システムであって、SANトポロジ・マップを生成するための第1の生成手段と、SAN問題判別情報テーブル(SPDIT)を生成するための第2の生成手段と、前記SANトポロジ・マップ及び前記SPDITを用い手SAN診断テーブル(SDT)を生成するための第3の生成手段とを、有することを特徴とするデータ処理システム。

(18) 前記SANトポロジ・マップはSANトポロジ・テーブルを有し、該SANトポロジ・テーブルの各々の行が通信アーキテクチャ要素(CAE)に対して一意的にマッピングされ、また前記SANトポロジ・テーブルの各々の列がCAEに対して一意的にマッピングされており、CAEはネットワーク・サービス・プロトコルを介して通信アーキテクチャ管理プログラム(CAM)によって首尾よく登録されたネットワーク接続装置であり、CAMは前記SANのための問題判別(PD)機能を有し、またSANのPD情報テーブル(SPDIT)を保持し、さらに前記通信アーキテクチャ(CA)は前記SPDITに格納された情報を伝達する能力を有するすべてのネットワーク接続要素を有することを特徴とする上記(17)に記載のデータ処理システム。

(19) 前記SPDITは、前記CA上の各プロダクト又は要素に対応付けられた少なくとも一つのデータ・レコードを有することを特徴とする上記(18)に記載のデータ処理システム。

(20) 前記CA上の各プロダクト又は要素に対応付けられた少なくとも一つのデータ・レコードは、プロダクト・ベンダ情報、プロダクト識別子情報、前記プロダクト又は要素によってサポートされた通信リンクの種類に関する情報、及び/又は前記プロダクト又は要素によって報告されるエラー情報の種類に関する情報からなる群から選択される1つ以上のデータ項目をさらに有することを特徴とする上記(19)に記載のデータ処理システム。

(21) 前記エラー情報の種類は、前記プロダクト又は要素が拡張リンク・サービス(ELS)登録リンク事象レコード(RLIR)をサポートするかどうかを示すことを特徴とする上記(20)に記載のデータ処理システム。

(22) 前記SDTは、前記CAMによってCAEから受け取ったエラーと前記SANトポロジ・マップからの情報とを格納することを特徴とする上記(20)に記載のデータ処理システム。

(23) ストレージ・エリア・ネットワーク(SAN)
50 内のエラーを処理するデータ処理システムであって、コンピュータ実行ステップとして、通信アーキテクチャ管

理プログラム（C A M）でのエラー・メッセージ受信する受信手段と、リアルタイム診断アルゴリズム（R D A）を用いて前記エラー・メッセージを処理する処理手段とを有し、前記C A Mは前記S A Nのため問題判別（P D）機能を有し、該C A Mは前記S A NのP D情報テーブル（S P D I T）を保持し、さらに前記C A Mによって管理された通信アーキテクチャ（C A）は前記S P D I Tに格納された情報を伝達する能力を有するすべてのネットワーク接続要素を有することを特徴とするデータ処理システム。

（2 4）前記C Aをサポートするネットワークは、バンド内ファイバ・チャンネル通信リンクとバンド外通信リンクとを有することを特徴とする上記（2 3）に記載のデータ処理システム。

（2 5）前記S A Nは、前記ネットワークに接続した複数のストレージ装置と、前記ネットワークに接続した複数のホスト・コンピュータとをさらに有し、前記複数のホスト・コンピュータの少なくとも一つはC A Mを有することを特徴とする上記（2 3）に記載のデータ処理システム。

（2 6）S A Nトポロジ・マップを生成する第1の生成手段と、前記S A Nトポロジ・マップ及び前記S P D I Tを用いてS A N診断テーブル（S D T）を生成する第2の生成手段とを、さらに有することを特徴とする上記（2 3）に記載のデータ処理システム。

（2 7）すでに前記C A Mによって受信され、かつ前記S D Tに格納された時間的に関連したエラー・メッセージを検索する一方で、時間的制約障害分離判別に対する時間的相関ウインドウ（T C W）の値を用いて前記受信エラー・メッセージを分析する第1の分析手段と、すでに前記C A Mによって受信され、かつ前記S D Tに格納された空間的に関連したエラー・メッセージを検索する一方で、空間的制約障害分離判別に対する空間的相関バス・データ構造（S C P）を用いて前記受信エラー・メッセージを分析する第2の分析手段と、をさらに有することを特徴とする上記（2 6）に記載のデータ処理システム。

（2 8）前記受信エラー・メッセージによって示されたエラーの種類に応じた重大度重みを用いて前記受信エラー・メッセージを分析する第3の分析手段を、さらに有することを特徴とする上記（2 7）に記載のデータ処理システム。

（2 9）ストレージ・エリア・ネットワーク（S A N）においてエラー情報を伝達するためのデータ処理システムで使用されるコンピュータ読み取り可能媒体のコンピュータ・プログラム・プロダクトであって、S A Nトポロジ・マップを生成する第1の命令と、S A N問題判別情報テーブル（S P D I T）を生成する第2の命令と、前記S A Nトポロジ・マップ及びS P D I Tを用いてS A N診断テーブル（S D T）を生成する第3の命令と、

を有することを特徴とするコンピュータ・プログラム・プロダクト。

（3 0）ストレージ・エリア・ネットワーク（S A N）においてエラー情報を伝達するためのデータ処理システムで使用されるコンピュータ読み取り可能媒体のコンピュータ・プログラム・プロダクトであって、通信アーキテクチャ管理プログラム（C A M）でのエラー・メッセージ受信する第1の命令と、リアルタイム診断アルゴリズム（R D A）を用いて前記エラー・メッセージを処理

- 05 する第2の命令とを有し、前記C A Mは前記S A Nのため問題判別（P D）機能を有し、該C A Mは前記S A NのP D情報テーブル（S P D I T）を保持し、さらに前記C A Mによって管理された通信アーキテクチャ（C A）は前記S P D I Tに格納された情報を伝達する能力
10 15 を有するすべてのネットワーク接続要素を有することを特徴とするコンピュータ・プログラム・プロダクト。
【図面の簡単な説明】
【図1】本発明が適用されるデータ処理システムの概略的構成を説明するための模式図である。

- 20 【図2】本発明が適用されるサーバ型データ処理システムの内部構成要素の一例を説明するためのブロック図である。
【図3】本発明の好ましい実施の形態にもとづいて実現されたS A N問題判別方法論に関するデータ処理システムの通信アーキテクチャを説明するための模式図である。

- 25 【図4】本発明の好ましい実施の形態にもとづくS A N問題判別情報テーブル（S P D I T）を説明するための表である。

- 30 【図5】S A Nの簡略化したネットワーク・トポロジを説明するためのブロック図である。

- 【図6】図5に示すS A Nのためのトポロジ・マップを提供する表である。

- 35 【図7】S A NのためのS A N診断テーブルを示す表である。

- 【図8】様々なエラーに対するリアルタイム診断分析で使用される重みを説明するための表である。

- 40 【図9】本発明の好ましい実施の形態にもとづいて单一障害S A N要素のS A N端末間障害分離用リアルタイム診断アルゴリズムのプロセスを説明するためのフローチャートである。

- 45 【図10】本発明の好ましい実施の形態にもとづいて单一障害S A N要素のS A N端末間障害分離用リアルタイム診断アルゴリズムのプロセスを説明するためのフローチャートである。

- 【図11】本発明の好ましい実施の形態にもとづいて单一障害S A N要素のS A N端末間障害分離用リアルタイム診断アルゴリズムのプロセスを説明するためのフローチャートである。

- 50 【図12】本発明の好ましい実施の形態にもとづいて單

ストレージ・エリア・ネットワーク (S A N) 内 でのエラー処理方法及びデータ処理システム

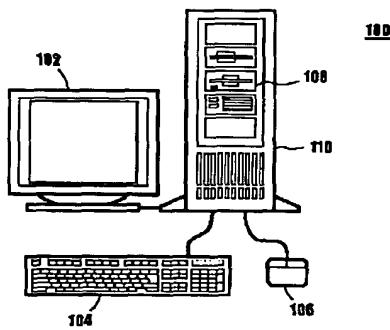
特開2001-249856

一障害SAN要素のSAN端末間障害分離用リアルタイム診断アルゴリズムのプロセスを説明するためのフローチャートである。

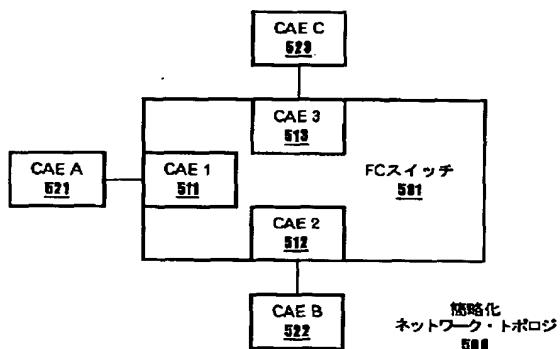
【符号の説明】

100	コンピュータ	232	I S A ブリッジ
102	ビデオ表示端末	233	N I O コントローラ
104	キーボード	234	I S A バス 233
106	マウス	235, 236	シリアル接続
109	ストレージ装置	05	フロッピー・ドライブ接続
200	データ処理システム	238	キー・ボード接続
202	プロセッサ	239	マウス接続
203	L 2 キャッシュ	240	不揮発性RAM (N V R A M)
205	6 XX バス	241	システム・ファームウェア
206	プロセッサ・カード	10	サービス・プロセッサ
207	プロセッサ	244	ネットワーク
208	L 2 キャッシュ	300	WindowsNT (登録商標) サーバ
210	システム・プレーナ	302	メインフレーム・コンピュータ
211	6 XX ブリッジ	304	Unix (登録商標) サーバ
212	メモリ・コントローラ	306	Linux (登録商標) サーバ
213	メモリ・カード	311, 312, 213	F C スイッチ
214	ローカル・メモリ	321, 322, 323	共有RA I D
215, 216	デュアル・オンライン・メモリ・モジュール (D I MM)	324	共有テープ
220, 221	P C I ブリッジ	325	共有テープ
222	システム・バス	20	通信リンク
223	I / O (N I O) プレーナ	341, 342, 343, 344, 345, 346, 3	
224	ネットワーク・アダプタ	47, 358, 349, 350, 351	バンド内
225, 226	カード・スロット	通信リンク	
227, 228	P C I バス	400	S P D I T
229	ハード・ディスク	25	ベンダ属性
230	S C S I ホスト・アダプタ	402	プロダクト識別子
231	グラフィック・アダプタ	403	情報の種類
		404	記述属性
		501	F C スイッチ
		30	FCスイッチポート
		521, 522, 523	C A E

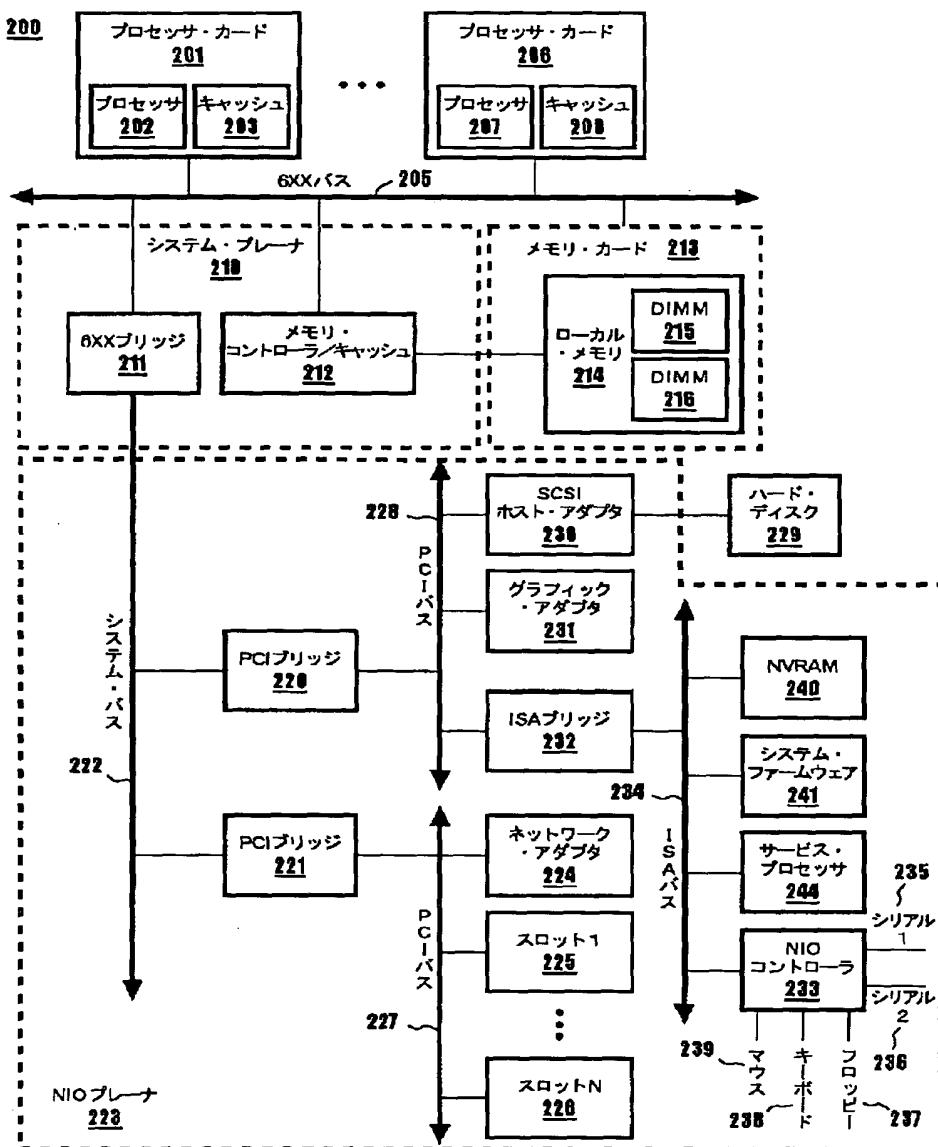
【図1】



【図5】

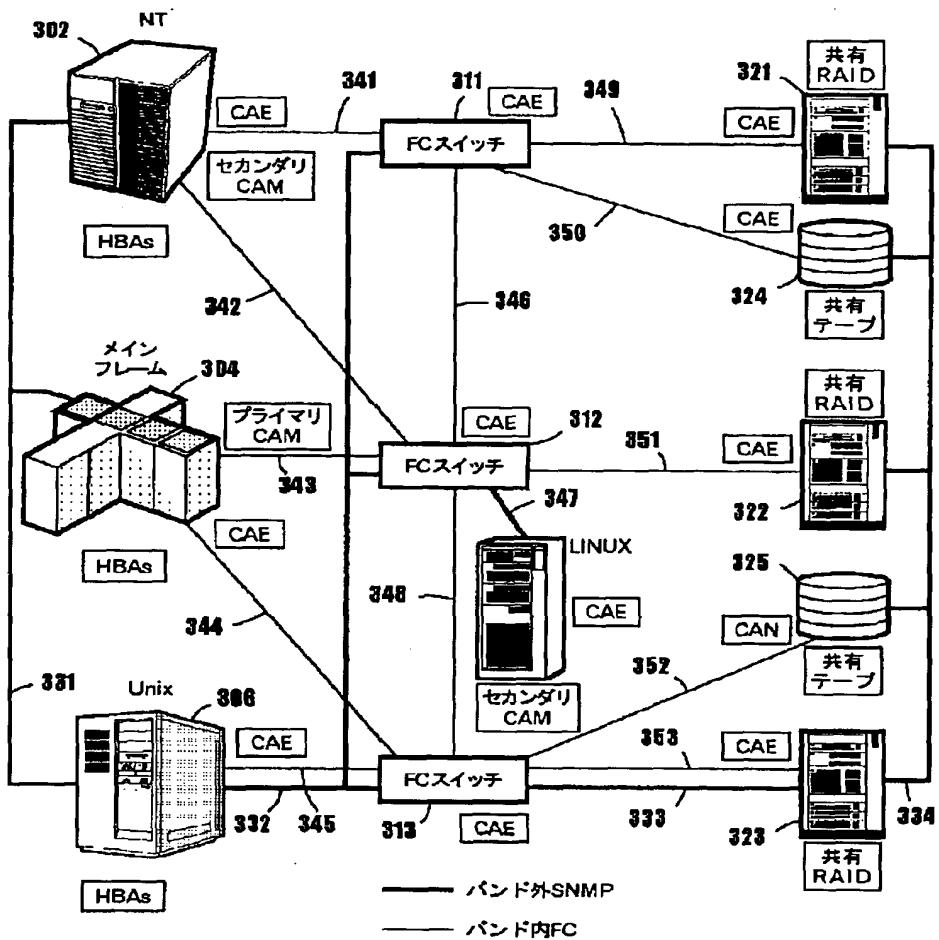


【図2】



ストレージ・エリア・ネットワーク (S A N) 内
でのエラー処理方法及びデータ処理システム

【図 3】



通信アーキテクチャを有するネットワーク

ストレージ・エリア・ネットワーク(SAN)内
でのエラー処理方法及びデータ処理システム

特開2001-249856

【図4】

SAN問題判別情報テーブル(SPDIT)

ベンダ <u>401</u>	プロダクト <u>402</u>	情報の種類 <u>403</u>	記述 <u>404</u>
FCスイッチCorp.	16 FポートFCスイッチ、 型123 プロダクトID=1	バンド内：FCリンク FC拡張リンク サービス： 登録リンク 事象レコード	ELS RLIR コンバチブル
⋮	⋮	⋮	⋮
FCストレージCorp.	4-ポート FC RAID 型234 プロダクトID=2	バンド外： SNMP MIB バンド内：FCリンク 拡張リンク RLIR	MIB=FC格納 ELS RLIR コンバチブル

【図6】

SANのためのトポロジ・マップ

	CAE A	CAE B	CAE C
CAE A	SPDIT/種類 情報 スイッチ・ポート CAE 1	(バスA→B)=(1→2) バス使用せず	(バスA→C)=(1→3)
CAE B	(バスB→A)=(2→1) バス使用せず	SPDIT/種類 情報 スイッチ・ポート CAE 2	(バスB→C)=(2→3)
CAE C	(バスC→A)=(3→1)	(バスC→B)=(3→2)	SPDIT/種類 情報 スイッチ・ポート CAE 3

ストレージ・エリア・ネットワーク(SAN)内
でのエラー処理方法及びデータ処理システム 特開2001-249856

【図7】

SAN診断テーブル

		CAE A	CAE B	CAE C
行	CAE A	CAE Aに 登録されたエラー リンク内リンク・タイムアウト 1/1/99@2.00.55		(パスA→C)=(1→3)
行2	CAE B		CAE Bに 登録されたエラー	(パスB→C)=(2→3)
行3	CAE C	(パスC→A)=(3→1)	(パスC→B)=(2→3)	CAE Cに 登録されたエラー リンク外コントローラ障害 1/1/99@2.00.01
行4	スイッチ CAE 1	CAE 1に 登録されたエラー		
行5	スイッチ CAE 2		CAE 2に 登録されたエラー リンク内リンク・エラー 1/2/99@4.55.09	
行6	スイッチ CAE 3			CAE 3に 登録されたエラー

【図8】

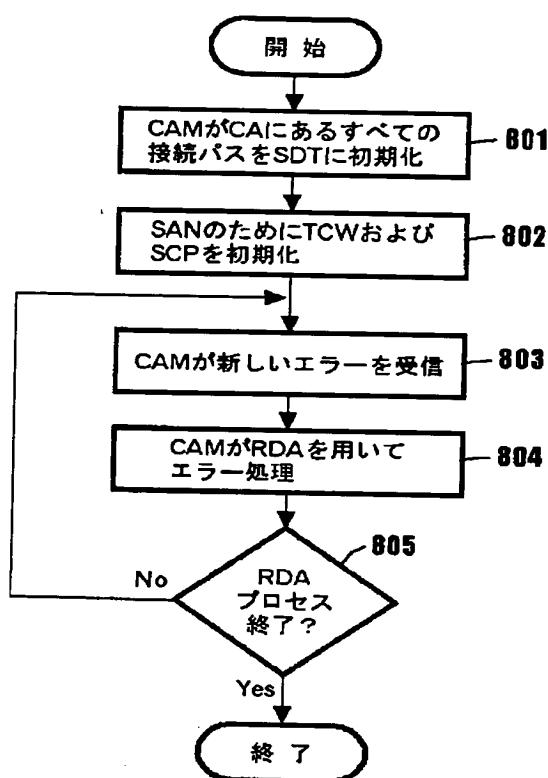
リアルタイム診断分析に対するエラー重み

装置エラーに対する重み	リンク・エラーに対する重み
コントローラ・ハードウェア障害 = H 装置ハードウェア障害 = H 装置ハードウェア・エラー回復可能 = H 装置タイムアウト = M コントローラ・タイムアウト = M 電源/ハードウェア障害 = H	リンク・ハードウェア不作動(スイッチ/N/L) = H リンク・ハードウェア障害回復可能 = M リンク・タイムアウト・エラー = L リンク・ハードウェア劣化 = H 電源/ハードウェア障害 = H

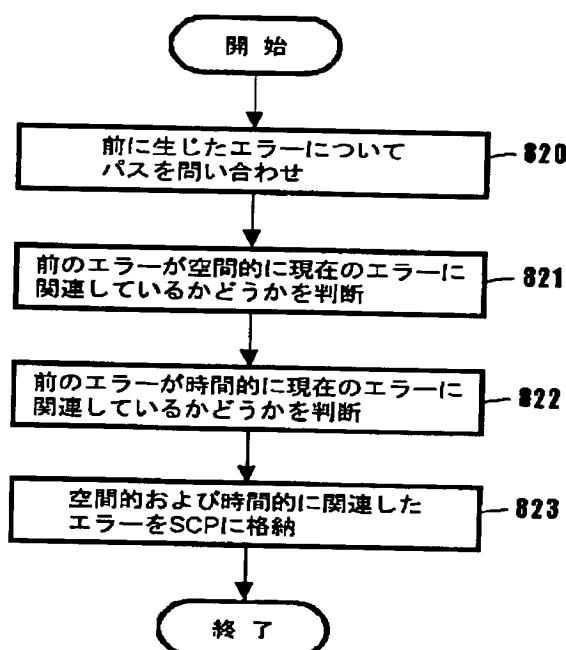
ストレージ・エリア・ネットワーク(SAN)内
でのエラー処理方法及びデータ処理システム

特開2001-249856

【図9】



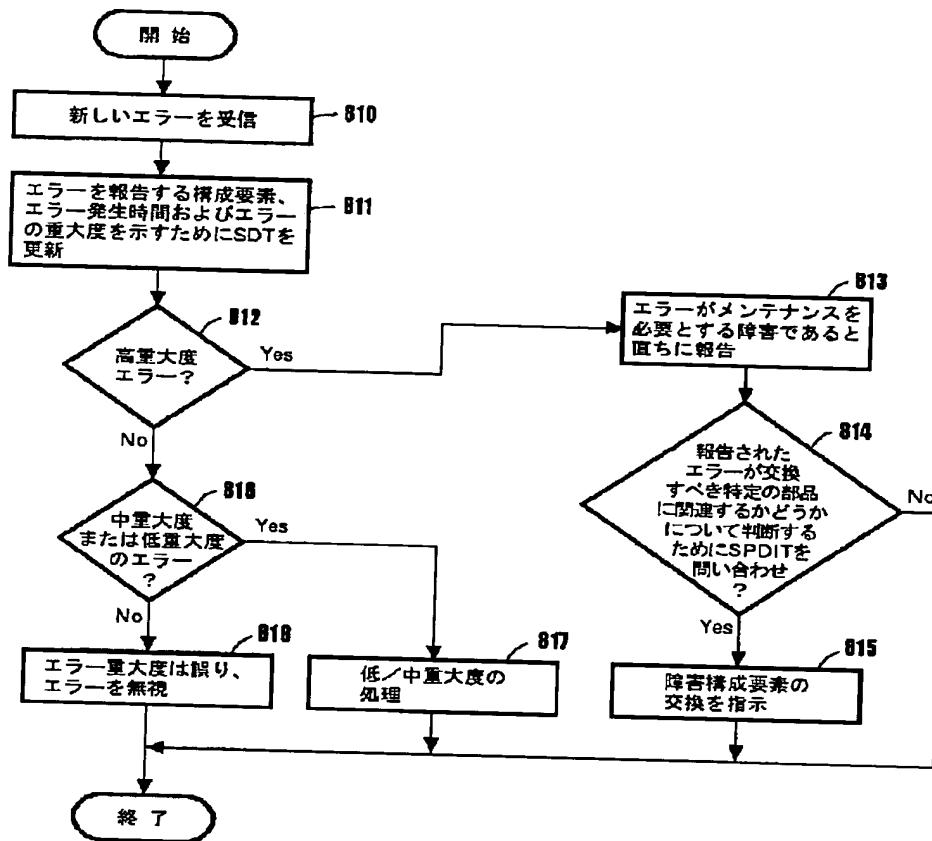
【図11】



ストレージ・エリア・ネットワーク(SAN)内
でのエラー処理方法及びデータ処理システム

特開2001-249856

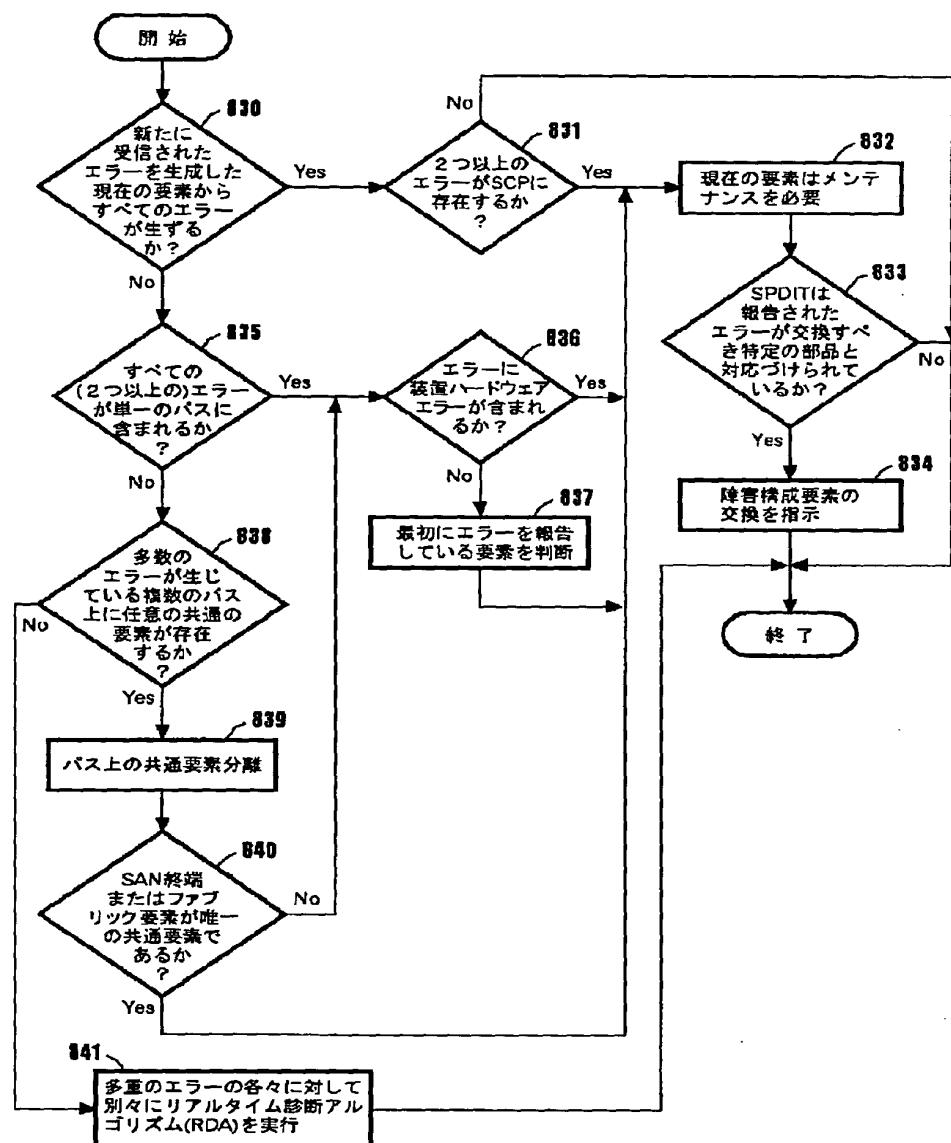
【図10】



ストレージ・エリア・ネットワーク (S A N) 内
でのエラー処理方法及びデータ処理システム

特開2001-249856

【図12】



フロントページの続き

(72)発明者 バリィ・スタンレイ・バーネット
アメリカ合衆国78759 テキサス州、オ
ースティン、シェイクスピリアン・ウェ
イ 11510

(72)発明者 ダグラス・クレイグ・ボッセン
アメリカ合衆国78731 テキサス州、オ
ースティン、ディア・リッジ・サークル
7801